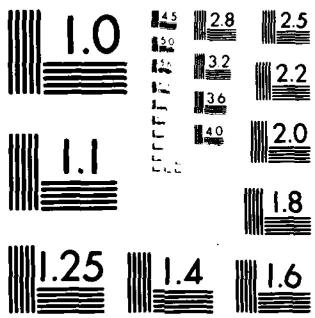
AD-A093 533 HAWAII UNIV HONOLULU DEPT OF ELECTRICAL ENGINEERING F/G 5/2
ADDRESSING AND DIRECTORY SYSTEMS FOR LARGE COMPUTER MAIL SYSTEM--ETC(U)
OCT 80 J J GARCIA-LUNA, F F KUO N00014-78-C-0498
UNCLASSIFIED CCIS-4 NL

1 of 1
AD A
093533

END
DATE
FILMED
2-81
DTIC



MICROCOPY RESOLUTION TEST CHART

12

AD A093533

Addressing and Directory Systems for
Large Computer Mail Systems
(Rep. No. CCIS-4)

LEVEL II

by

Jose J. Garcia-Luna
Frank F. Kuo
Dept. of Electrical Engineering
University of Hawaii
Honolulu, HI 96822

October 1980

The efforts of one of the authors (Kuo) on this paper was supported by the Office of Naval Research under Contract N00014-78-C-0498.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
14 CCIS-4	AD-A093 533	

4. TITLE (and Subtitle)

6 Addressing and Directory Systems For Large Computer Mail Systems

5. TYPE OF REPORT & PERIOD COVERED

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)

10 Jose J. Garcia-Luna
Franklin F. Kuo

8. CONTRACT OR GRANT NUMBER(s)

15 N00014-C-78-0498

9. PERFORMING ORGANIZATION NAME AND ADDRESS

Department of Electrical Engineering
University of Hawaii at Manoa
Honolulu, HI 96822

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS

Office of Naval Research
Code 437
Arlington, VA 22217

12. REPORT DATE

11 October 1980

13. NUMBER OF PAGES

19

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

12 21

15. SECURITY CLASS. (of this report)

Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Electronic Mail
Computer Message Systems
Computer Mail

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper, we present a general description of the addressing schemes necessary to provide efficient identification and delivery services in large computer mail systems. The structure of the required directory system is described in detail under the assumption of the existence of a communication protocol utilized to communicate the delivery processes. For the purposes of completeness, the basic structure of an architecture for large computer mail systems is introduced at the beginning of this work.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

400634

Addressing and Directory Systems for
Large Computer Mail Systems

by

Jose J. Garcia-Luna
Franklin F. Kuo

ABSTRACT

Computer mail systems constitute a family of services aimed at the enhancement of human communication by providing rapid information transmission and interactive computer tools to process information before and after transmission. This new application of computer networking presents various problems yet to be solved and due to a diversity of factors which range from user population size to complexity of market and diversity of user's service needs. To the present, various models for computer mail systems have been proposed and a good number of systems have been implemented. Yet, it appears that the techniques used to build the current systems are suitable neither for larger systems nor for system interconnection, and the models proposed so far do not provide answers to the problems which arise when large computer mail systems are to be implemented or various computer mail systems have to be interconnected. In this paper, we present a general description of the addressing schemes necessary to provide efficient identification and delivery services in large computer mail systems. The structure of the required directory system is described in detail under the assumption of the existence of a communication protocol utilized to communicate the delivery processes. For the purposes of completeness, the basic structure of an architecture for large computer mail systems is introduced at the beginning of this work.

INTRODUCTION

There are two main functions any computer mail system (CMS) must provide: message creation and retrieval, and message delivery. The simplest computer mail systems only provide for the delivery of the messages between users' mailboxes, and more sophisticated systems provide, in addition, interactive computer tools for message processing.

We will focus on large computer mail systems and unions of computer mail systems and describe an architectural model which specifies the user functions of such systems. For the specification of the proposed architecture we rely on the assumption that an underlying "implementation level" provides the functions taken for granted by the architecture, such an implementation level is made up of those processes which provide the transport services that enable the various components of the CMS to communicate with each other in order to provide the services expected by the users of the system.

Our architecture is defined at the application level and is formed by a set of independent functional components, each of which supports a specific set of functions to be provided. The functional components to be defined constitute processes which should not be directly related with underlying implementation considerations such as host computers, intelligent terminals, or communication networks.

We define three functional components in a computer mail system: mailboxes, mailers, and gateway mailers.

MBX MAILBOX is the process dedicated to the creation and retrieval functions of the system; this is also the interface between user and delivery service, that is, all user's messages are created by the interaction between the user and his mailbox. A mailbox is formed by three elements, namely:

1. The user interface

2. The mailbox database
3. The mailer interface

A personal directory with system addresses of recipients is maintained in the mailbox database together with any other information useful for the user, such as previous incoming and outgoing messages. The user interface is the process by which the user creates and retrieves messages. The mailer interface is the process by which the mailbox communicates with its local mailer.

MLR MAILER is the process that manages the delivery of messages to and from mailboxes and communicates when necessary with other mailers or the gateway mailer of the mailing network. Each mailer serves a set of one or more mailboxes, called local mailboxes, and each mailbox communicates with one and only one mailer, called the local mailer. A mailer is formed by four processes:

1. The mailbox interface which communicates with the local mailboxes
2. The network interface which communicates with other mailers
3. The mailing protocol handler which is in charge of the delivery of messages to and from the local mailboxes
4. The mailer's directory database which contains addressing information as well as control information about the negotiations handled by the mailer

GMR GATEWAY MAILER is the process that manages the delivery of messages between a mailing network and other mailing networks connected to it. Each gateway mailer serves a unique set of local mailers and each local mailer communicates with one and only one gateway mailer, called the local gateway mailer. This component performs all the necessary transformations of format, addressing and data representation to communicate the local mailing

network with others and with message systems such as Telex or TWX. The structure of a gateway is basically the same than that of the mailer, namely it contains:

1. The gateway directory database
2. The mailer interface
3. The system interface
4. The mailing protocol handler

The primary role of a gateway mailer at the intrasystem level is to serve as an intermediate hop between source and destination mailers to alleviate the identification tasks of each mailer in a network; at the intersystem level, its role is to support and enforce the independence and cooperation of intersystem mailing processes.

MNT--MAILING NETWORK is the union of logically connected mailers (implying both reachability and addressability). The topology of an MNT is completely independent of the underlying communication network's topology. The delivery of a message within an MNT is called <u>intranet</u> delivery. When more than one MNT is involved, the process is called <u>internet</u> delivery. Each MNT has associated with it a gateway mailer (Fig. 1). An MNT is related with the server of the computer mail system and with geographical qualifiers of the mailers being grouped together, e.g., the set of one or more states, or one or more countries, that the mailers serve.

CMS--COMPUTER MAIL SYSTEM is the union of logically connected MNTs offered by the same server, and the corresponding mailboxes (Fig. 2).

UCMS--UNION OF COMPUTER MAIL SYSTEMS is the union of cooperative CMSs, formed by the interconnection of the MNTs of the various servers (Fig. 3).

ADDRESSING

Users compose messages to be delivered to other users, groups of users, organizations, places, processes, machines, etc., which we call recipients. Recipients are viewed by the users as a set of named objects. Unfortunately, users do not name the same recipient in a unique form. A normal user usually describes a recipient in terms of who he is, where he is, what he does, and what relations he has with other entities known by the server.

In the proposed architecture, users are allowed to describe the recipients in terms of addresses much like that of the U.S. Postal System and not their system addresses, thus the users are freed from the task of knowing the addressing structure of the system. A common format must be followed by the users when they describe the recipients so that the system can understand the information provided by the users. Such a format must be compatible with the common forms in which users describe the recipients, and according to this we define a user's naming standard that we call the NOLS address [3]. A NOLS address is a set of four flexible fields which completely describe a recipient in a user-oriented, machine-processable form. The four fields of a NOLS address are described as follows:

N-field --- Consists of the name and/or role of the recipient.

O-field --- Consists of the information about the organization related to the recipient.

L-field --- Consists of the main information about the organization's geographical location.

S-field --- Consists of the name of the system which offers the computer mail services to the recipient.

Each field of a NOLS address corresponds to a partition of the total population of recipients. When the four fields are correctly and unambiguously

specified, the ultimate partition uniquely corresponds to the intended recipient. Since users are allowed to ask for the delivery of messages for any amount of information they provide in the NOLS addresses, these addresses may in some cases be ambiguous, lack information in a certain field, or contain redundant information, and responsibility of the server is to use the information contained in the NOLS addresses in the best possible way to try to deliver the messages to the desired recipients--and only to the desired recipients.

By organization we mean any entity well-known by the server and which corresponds to a partition of the total community of users; thus, an organization can be: a company, a regional center, a regional computer mail server, a branch office, etc. The contents permitted in each field of a NOLS address depends on the sophistication of the system's identification database.

Even though users compose the messages to be delivered to recipients, the actual delivery of the messages is not directed to the recipients but to some logical entities to which the server is prepared to deliver the messages. Those entities consitute the processes we call mailboxes, and in the proposed architecture the population of mailboxes in the system is viewed as a set of uniquely named objects. The responsibility of the system is to map a NOLS address describing a recipient into a system address referring to the mailbox directly related with the recipient.

For purposes of message delivery, the community of mailboxes is partitioned into various subsets following a hierarchical structure which may reflect the structure of the community served by the system.

We define the term <u>NEXUS</u> to be an association established between two functional components for the purpose of message delivery [3]. A

NEXUS address is the identification of a NEXUS in the system according to the system's structure. An end-to-end NEXUS associates the originating mailbox with the destination mailbox.

A NEXUS address consists of two parts, an origination part and a destination part, which uniquely identify the originating and destination processes respectively. A mailbox is identified in a NEXUS address by specifying:

1. The name of the local mailing network;
2. The name of the local mailer; and
3. The name of the mailbox.

That is:

End-to-End NEXUS address = [(OMNT,OMLR,OMBX),(DMNT,DMLR,DMBX)]

where the prefix O means originating and the prefix D means destination.

Thus, the NOLS address is a user-oriented naming standard with which a sender describes the recipients, and the NEXUS address is a system-oriented address which the system identifies the mailboxes related with the recipients. In our architecture, the users are completely unaware of the system's addressing structure, and therefore the responsibility of the system is to map a NOLS address into a NEXUS address for the purpose of message delivery.

Generally speaking, message delivery in our architecture consists of two phases: the NEXUS establishment phase and the user message delivery phase. In the NEXUS establishment phase the system identifies the recipient according to the NOLS address provided by the sender (that is, it obtains the end-to-end NEXUS address). Once the system knows the address of the recipient's mailbox, the delivery of the user message takes place.

In large computer mail systems and in a union of computer mail systems, independence and distribution of functions are of primary importance to allow for system extensibility. In the proposed architecture, each functional component is dedicated to serve a unique set of recipients associated with a partition of the mailbox community and has complete knowledge of only a small fraction of the total system. Key points to note are:

1. A mailbox knows only about the existence of its local mailer,
2. A mailer knows only about the existence of its local mailboxes, other mailers in the same mailing network and the local gateway mailer, and
3. A gateway mailer knows only about the existence of the local mailers and other gateway mailers in the total system.

Because of this partial knowledge of the system by its parts, an addressing identification algorithm must be provided for the components to obtain the end-to-end NEXUS address from a NOLS address. Such an algorithm requires both a hierarchical organization of the various elements of the system, and the establishment of a set of rules with which those elements interact with each other to exchange information. The organization of the components of the system involves the specification of three main issues:

1. The relation each functional component has with the community served by the system,
2. The distribution of the identification information needed to map NOLS addresses into NEXUS addresses, and
3. The specification of the functions each functional component has concerning recipient identification.

These issues constitute the specification of the directory system of the architecture to be described in the following section.

DIRECTORY SYSTEM

A directory is defined to be a listing of identification information about recipients or subsets of the community of recipients. A directory system is the union of all the directories in a CMS, that is, a system-wide identification database. In our architecture, the directory system is formed by the union of the directories in mailboxes, mailers and gateway mailers; therefore, such a system constitutes a distributed identification database.

The primary purpose of the directory system in a CMS is to allow for the reliable and efficient identification of the recipients of the messages being delivered. By identification of a recipient we mean the identification of his system address from a NOLS address given by the sender, and in our architecture this identification service is provided in an automatic form when users instruct the system to deliver the messages to the recipients.

Another use of the directory system is for information retrieval. Such a service allows users to obtain identification information about recipients, that is, who a recipient is, what he does, where he is, etc.

There are four main issues in the design of a directory system:

1. What information is maintained in the directory database,
2. How the information is distributed in the system,
3. How the information is maintained or updated, and
4. What procedure is followed to use the information.

The directory system must maintain the identification information of all the recipients in the computer mail system. Different information must be stored for different types of recipients [2], and the amount of information to be maintained about each entry depends on the sophistication of the computer mail system. In large systems and unions of systems we can expect to have:

1. Diversity of data elements to be stored according to the type of recipient.

2. Diversity in the quality of information stored from one system to another.

3. Inconsistency of information (i.e., differences in format) stored in different systems and even in different mailers in the same system.

4. Large amounts of information to be stored to allow more flexibility in the NOLS addresses and provide for a better information retrieval service.

Due to the amount of information to be maintained in a computer mail system intended to serve large communities, the existing storage constraints, the desired fast response times, and the necessary independence between systems, an appropriate form to distribute the identification information among the various functional components must be specified in the architecture. We define the term complete identification information to be that information which completely specifies who the recipient is, where he is, what he does, and who belongs to the recipient (in the case of groups and organizations).

A complete identification information obviously contains the data required in a NOLS address and a NEXUS address. The flexibility allowed in a NOLS address depends on the quality of the identification information maintained in the directory system about a recipient.

We propose a hierarchical directory system [3] to maintain the identification information in large computer mail systems and unions of computer mail systems. In such a directory system, the directories in one hierarchical level serve as the means for the directories in the level below to interact with each other in an efficient form. A directory queries a set

of one or more directories in other mailers or gateway mailers based on the NEXUS address(es) it obtains from the NOLS address(es) provided by the sender. Fig. 4 shows the distribution of identification information among the various components of the architecture.

Each mailbox maintains a personal directory containing information about the recipients frequently addressed by the user (i.e., the end-to-end NEXUS address plus some other information). A personal directory can be updated when an entry is not found upon a user's request for message delivery or when a forwarding action due to a change of address had to be taken when delivering a user's message. The union of personal directories is called the user directory system (UDS in Fig. 4).

Each mailer maintains three address directories: a local directory with complete identification information about all the local recipients, a network directory with mailer system addresses (including the gateway mailer system address) according to the relations organizations have with the mailers in the system and a forwarding directory with the system addresses of those recipients who moved to other mailers. Each gateway mailer maintains three address directories: a network directory with the system addresses of the local mailers according with their relations with groups and organizations, a system directory with a gateway mailer system address for various of the mailing networks in the system and a forwarding directory with the system address of those groups and organizations who moved to another mailing network.

The union of local and network directories in a mailing network is called a network directory system (NDS in Fig. 4), where each local directory maintains complete identification information about local recipients and the network directories maintain pointers to each other, according to the relations between organizations and mailers, to optimize the interaction between

local directories to resolve queries about remote recipients. In such a system, queries about local recipients are directly resolved by the local directories whereas queries about remote recipients are referred to a set of one or more remote network directories. Due to the information obtained from the NOLS addresses, on the average, only a small number of remote network directories will have to be queried.

Since the network directories only contain mailer system addresses about the groups and organizations in the MNT and pointers to the GMR for the case of groups and organizations in other MNT, the updating requirements for a network directory system are very low.

The union of system directories is called system directory system (SDS in Fig. 4). Such a system constitutes the means for network directories in different mailing networks to communicate with each other in order to resolve user queries concerning recipients in remote MNTs. The updating requirements of a system directory are very low because the various system directories only maintain pointers to each other.

No restriction is imposed on the logical connections between network directories and between system directories established by means of the pointers they maintain to each other, in other words, the topology of an NDS or an SDS is arbitrary.

The information maintained in the local directories about the local recipients needs no updating. The information maintained in the network directories and system directories however does need to be updated since the pointers (system addresses) in these directories must be updated according to the changes in the organization addresses. Since these pointers are associated with groups and organizations, they rarely change, and because of this, the updating strategy can be simplified as follows: A command asking for the identification of a recipient or the delivery

of a message is received in a mailer or gateway mailer. The queried mailer or gateway mailer then determines whether or not the organization specified in the NOLS address is listed in its directories and whether it has moved to another mailer. If it has moved, an UPDATE command is sent back to the sending mailer or gateway mailer.

The general algorithm to map a NOLS address into an end-to-end NEXUS address is shown in Fig. 5. Such an algorithm assumes the existence of:

1. A communication protocol which specifies how the various mailers and gateway mailers communicate, and
2. A well defined storage format understood by mailers and gateway mailers.

The communication protocol will be described in a forthcoming paper.

CONCLUSIONS

In this paper we have presented a general description of the necessary addressing standards in large computer mail systems with transparent addressing services. First, the basic structure of an architecture for large computer mail systems and unions of computer mail systems was outlined. A hierarchical directory system was proposed to organize the interaction of the various components of the architecture for the purposes of message delivery and recipient identification. Such a system was defined over the architecture previously specified and assumed the existence of:

1. A communication protocol that specifies how the components of the architecture communicate with each other, and
2. A storage format well-known by each one of the delivery processes.

The specification of such communication and storage protocol constitutes in itself a very important design problem related to computer mail systems, Finally, it should be emphasized that the standards presented in this paper constitute a starting point in the development of a complete addressing scheme for large systems. The communication protocol necessary for the communication of the delivery processes in computer mail systems will be described in a forthcoming paper.

REFERENCES

[1] Chu, W. W. "Performance of File Directory Systems for Databases in Star and Distributed Networks," AFIPS Proceedings, Vol. 45, 1976, pp. 577-587.

[2] Feinler, E. J. "The Identification Database in a Networking Environment," NTC 1977 Conference Record, 1977, pp. 21:3-1 - 21:3-5.

[3] Garcia-Luna, J. J. "A Study of Computer Mail Services," Master's Thesis, University of Hawaii at Manoa, August 1980.

[4] Henderson, D. A., Jr. and Myer, T. H. "Issues in Message Technology," Proceedings of the Fifth ACM/IEEE Data Communication Symposium, 1977, pp. 6-1 - 6-9.

[5] Panko, R. R. "The Cost of Computer Mail and Other Media," Working Paper, Project 6859, SRI International, 333 Ravenswood Avenue, Menlo Park, CA 94025, July 1978.

[6] Pickens, J. R., Feinler, E. J., & Mathis, J. E. "The NIC Name Server -- A Datagram-Based Information Utility," Proceedings of the Fourth Berkeley Conference on Distributed Data Management and Computer Networks, August 1979.

[7] Shock, J. F. "Internetwork Naming, Addressing and Routing," IEEE 1978 COMPCON Fall, pp. 72-79.

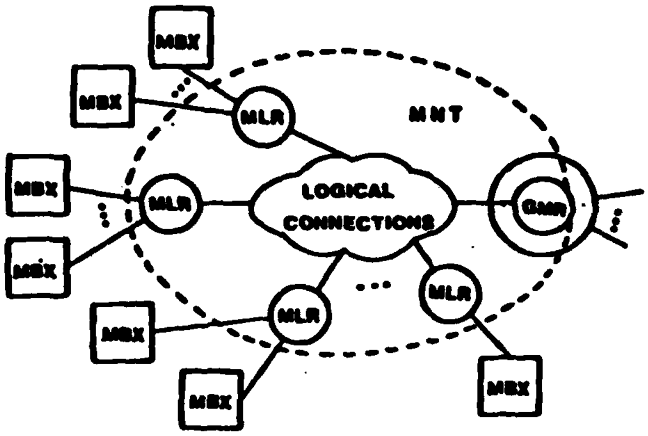


Fig. 1 Mailing Network

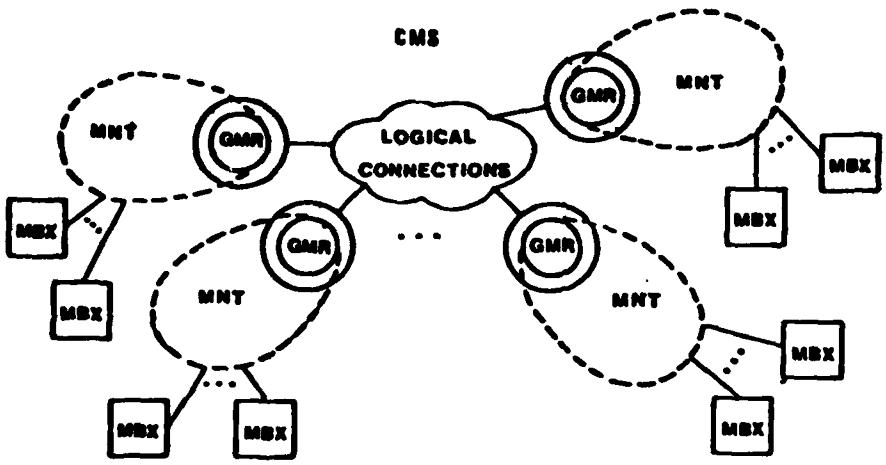


Fig. 2 Computer Mail System

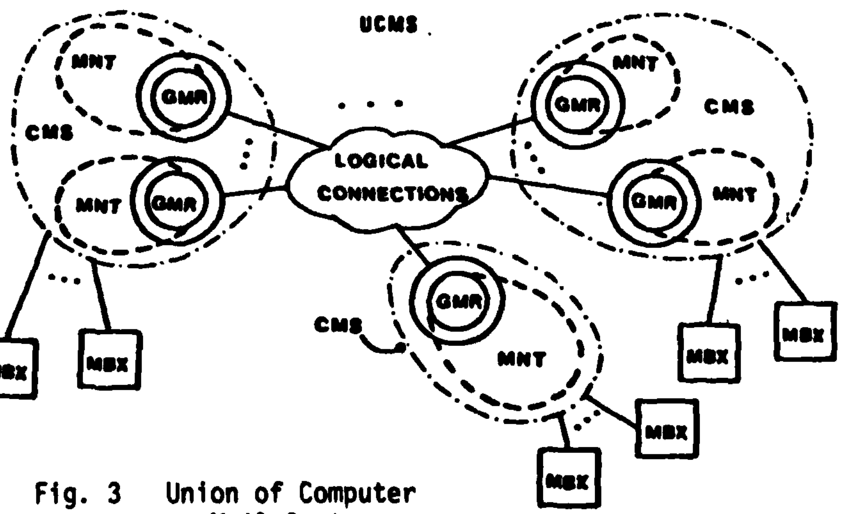


Fig. 3 Union of Computer Mail Systems

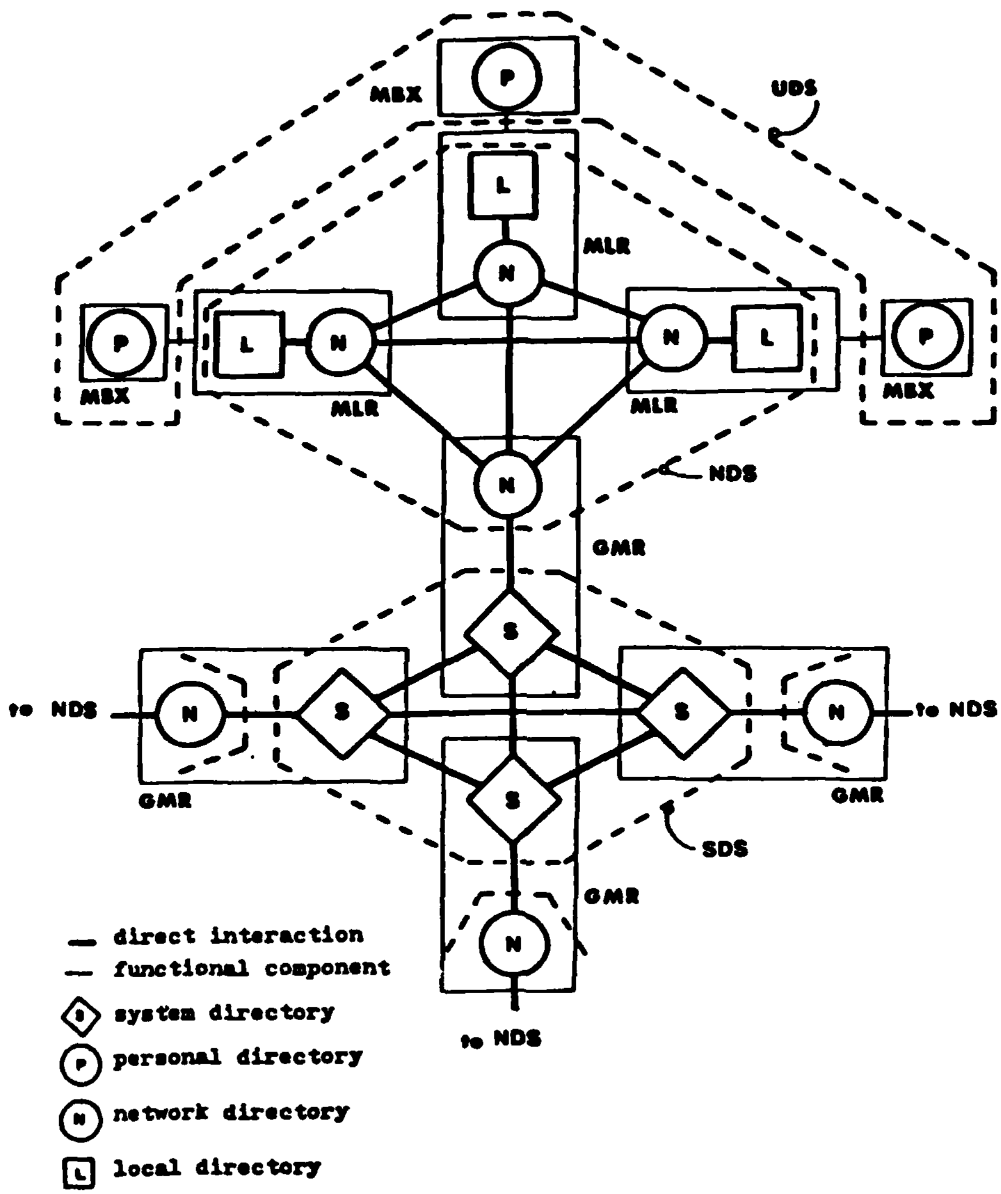


Fig. 4 Directory System

NOLS-TO-NEXUS MAPPING ALGORITHM

MAILER ACTIONS.

Fig. 5a

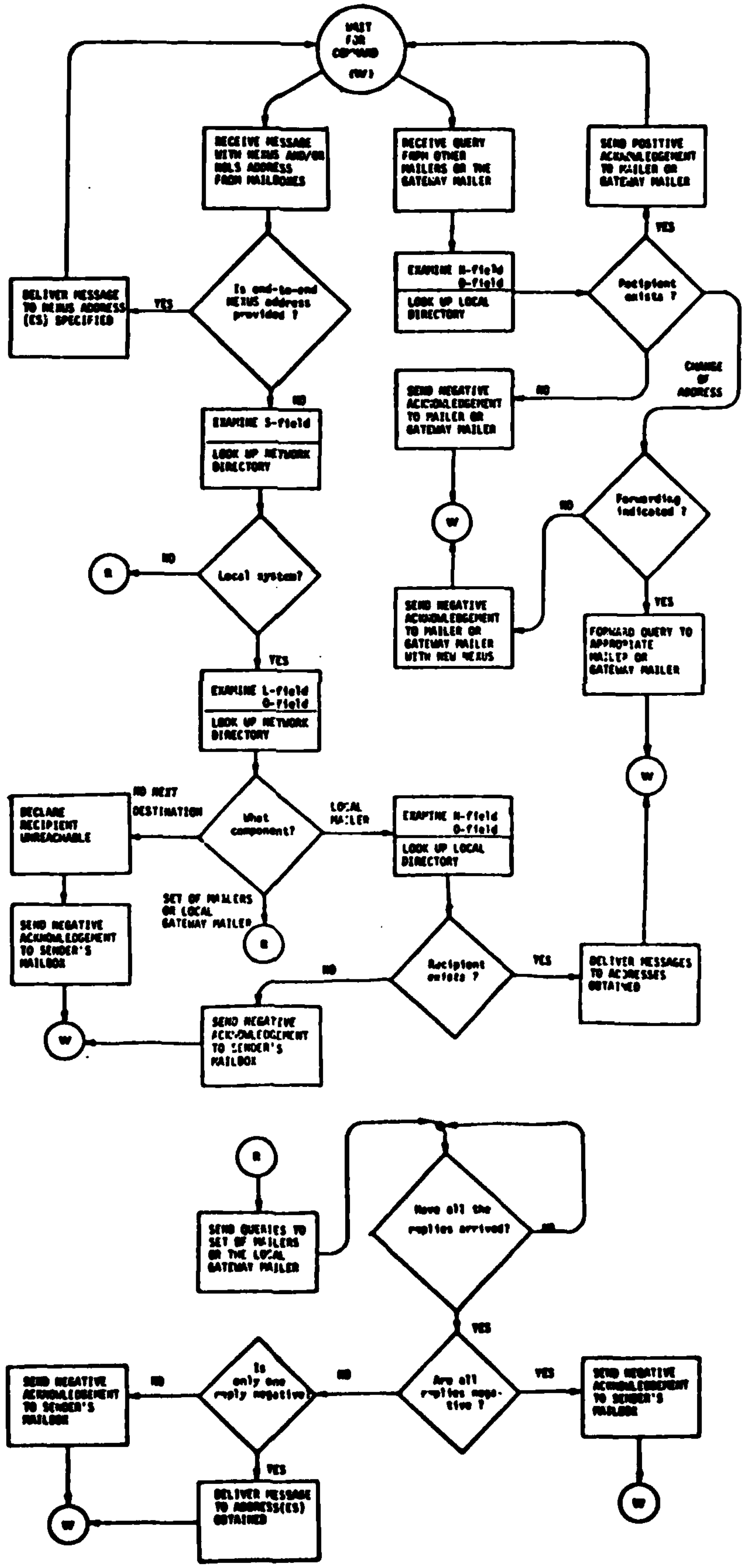


Fig. 5b ACTIONS TAKEN BY THE GATEWAY MAILER.